

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-326286
 (43)Date of publication of application : 08. 12. 1998

(51)Int. Cl. : G06F 17/30
 G06T 1/00
 G06T 7/00

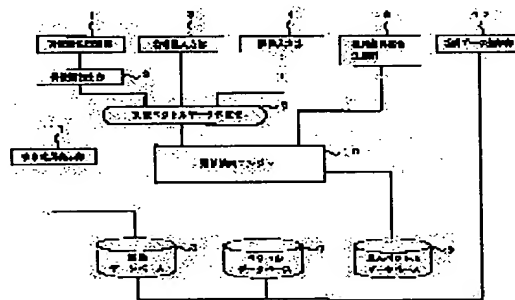
(21)Application number : 09-136816 (71)Applicant : MITSUBISHI ELECTRIC CORP
 (22)Date of filing : 27. 05. 1997 (72)Inventor : TAKAHASHI MASANOBU

(54) SIMILARITY RETRIEVAL DEVICE AND RECORDING MEDIUM WHERE SIMILARITY RETRIVAL PROGRAM IS RECORDED

(57)Abstract:

PROBLEM TO BE SOLVED: To improve the precision of similarity retrieval and to reduce the possibility that important similar data is omitted from retrieval results by retrieving vector data similar to object vector data and displaying retrieval results for individual retrieval conditions.

SOLUTION: A retrieval condition set generation part 6 generates a set of retrieval conditions for similarity retrieval. Vector data with high similarity to an object vector data which meet individual retrieval conditions belonging to the retrieval condition set are retrieved by a similarity retrieval engine 10 from a vector data base 7 by using weight vectors, corresponding to the retrieval conditions, among weight vectors in a weight vector data base 9. Then a retrieval result display part 11 displays some of attributes of retrieved similar vector data and images corresponding to the similar vector data among images in an image data base 8 for individual retrieval conditions.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-326286

(43) 公開日 平成10年(1998)12月8日

(51) Int.Cl.⁹

識別記号

F I

G 0 6 F 17/30

G 0 6 F 15/403

3 5 0 C

G 0 6 T 1/00

15/40

3 7 0 B

7/00

15/62

3 3 0 G

3 9 0 Z

15/70

4 6 0 B

審査請求 未請求 請求項の数 8 O L (全 13 頁)

(21) 出願番号

特願平9-136816

(22) 出願日

平成9年(1997)5月27日

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 高橋 正信

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

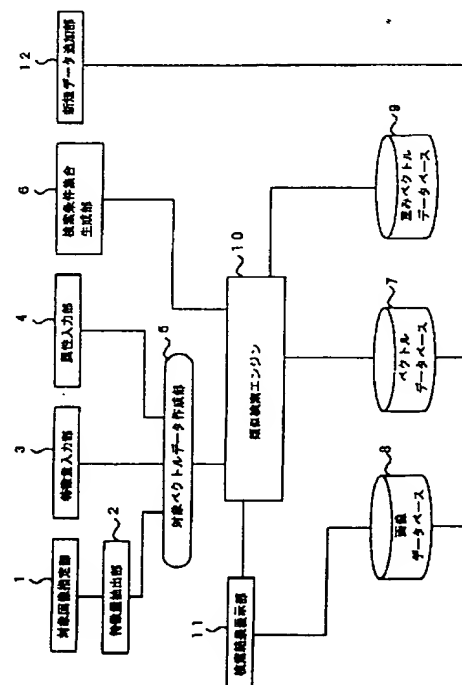
(74) 代理人 弁理士 宮田 金雄 (外2名)

(54) 【発明の名称】 類似検索装置及び類似検索プログラムを記録した記録媒体

(57) 【要約】

【課題】 類似検索の精度が向上し、また重要な類似データが検索結果からぬけおちる可能性が低い類似検索装置を提供する。

【解決手段】 検索条件集合生成部6を備え、検索条件集合生成部6で生成された個々の検索条件ごとに、検索条件に対して最適化された重みベクトルを用いて対象ベクトルデータに対する類似ベクトルデータを検索し表示する。



【特許請求の範囲】

【請求項1】 複数の対象に対して各々作成され、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象に対するベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果を個々の検索条件ごとに表示する検索結果表示部とを備えたことを特徴とする類似検索装置。

【請求項2】 複数の対象に対応する複数の対象画像を蓄積する画像データベースと、上記複数の対象に対して各々作成され、上記対象画像の特徴を定量的に表す特徴量と、上記対象を特徴づける属性とをベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象の画像と属性を基にベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果を対応する画像と共に個々の検索条件ごとに表示する検索結果表示部とを備えたことを特徴とする類似検索装置。

【請求項3】 画像の特徴量を抽出する特徴量抽出部と画像の特徴量を入力する特徴量入力部との少なくとも一方を備え、画像の特徴を定量的に表す特徴量として、上記特徴量抽出部において抽出された特徴量と上記特徴量入力部より入力された特徴量との少なくとも一方を用いることを特徴とする請求項2記載の類似検索装置。

【請求項4】 画像として生体組織の画像を用い、検索条件として病理組織診断における診断名を用いることを特徴とする請求項2または3記載の類似検索装置。

【請求項5】 画像の特徴量として、核領域の数、核領域の面積、核領域の形状、核領域の円形度、核領域の色、核領域の色度、空隙領域の数、空隙領域の面積、空隙領域の形状、空隙領域の円形度、間質領域の数、間質領域の面積、間質領域の形状、間質領域の円形度、間質領域の色、間質領域の色度、管空領域の数、管空領域の面積、管空領域の形状、管空領域の円形度、画像のテクスチャ、ウェーブレット変換値を用いて算出される特徴量、上皮細胞が筋上皮細胞を伴う2層構造を示す程度、線維化の程度、乳頭状パターンの有無、ふるい状パターンの有無、壊死物質の有無、充実性パターンの有無、画

像の色、及び色度を用いて算出される特徴量の少なくとも一つを用いることを特徴とする請求項4記載の類似検索装置。

【請求項6】 複数の検索条件にそれぞれ対応した複数の重みベクトルを蓄積する重みベクトルデータベースを備え、類似検索エンジンが用いる重みベクトルが、検索条件ごとに特定されていることを特徴とする請求項1ないし5のいずれかに記載の類似検索装置。

【請求項7】 類似検索対象に対して、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに出力する検索結果出力手順とをコンピュータに実行させる類似検索プログラムを記録した記録媒体。

【請求項8】 類似検索対象の画像と上記対象を特徴づける属性を基に、対象画像の特徴を定量的に表す特徴量と上記対象の属性とをベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに対応する画像と共に出力する検索結果出力手順とをコンピュータに実行させる類似検索プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、蓄積されたデータ群の中から、指定された検索対象に対するデータに類似するデータを検索する類似検索装置、及びこのような類似検索プログラムを記録した記録媒体に関するものである。

【0002】

【従来の技術】従来の類似検索装置として、例えば雑誌「Incremental Instance-based Learning of Independent and Graded Concept Descriptions」D. Aha, Proceedings of the Sixth International Workshop on Machine Learning, 1987、及び雑誌「A Nearest Hyperrectangle Learning Method」S. Salzberg, Machine Learning, 6, pp. 251-276, 1991に示されたものがある。図8はこのよ

うな従来の類似検索装置の一例を示すブロック図である。

【0003】図において、1は検索対象の画像を指定する対象画像指定部、2は対象画像指定部1により指定された対象画像の特徴を定量的に表す特徴量を抽出する特徴量抽出部、4は対象画像に関する特徴量以外の、対象を特徴づける属性を入力する属性入力部、5は特徴量抽出部2により抽出された特徴量と属性入力部4により入力された属性とをベクトル構成要素とする対象ベクトルデータを作成する対象ベクトルデータ作成部、7は上記と同様、特徴量と属性とをベクトル構成要素とするベクトルデータにより、複数の対象に対して作成された複数のベクトルデータをデータベースとして蓄えたベクトルデータベース、8は上記複数の対象に対応する複数の画像をデータベースとして蓄えた画像データベース、13は類似検索エンジンにおいてデータ間の類似度を算出するために用いるベクトルデータの各構成要素につけられた重みを構成要素とする重みベクトル、10はベクトルデータベース7に蓄えられた複数のベクトルデータの中から、対象ベクトルデータ作成部5で作成された対象ベクトルデータに類似したベクトルデータを探し出す類似検索エンジン、11は類似検索エンジン10により検索された類似するベクトルデータ、およびこのベクトルデータに対応する画像を表示する検索結果表示部、14は類似検索エンジン10が示した類似するベクトルデータおよび画像が利用者にとって正解であるかどうかを判定する回答指定部、15は回答指定部14において判定された結果をもとに、重みベクトル13を更新する重みベクトル更新部、12は新たにベクトルデータと画像をそれぞれベクトルデータベース7と画像データベース8に登録する新規データ追加部である。

【0004】次に動作について説明する。例えば電子カルテや医療画像データベースなどの医療情報、設計図などの設計情報などにおいては、発生するデータを蓄積し、その中から新しい目的に見合ったデータを選び出して参照する場合に、データをベクトルデータの形に整理してデータベースに蓄え、新しい目的を表現した検索対象のベクトルデータとデータベース中の各データとの類

似度を算出して、データベースの中で最も類似するデータを参照するという、ベクトルデータの類似検索技術が応用されている。ここでは、一例として、病理組織診断の支援を目的として、検査用の病理組織画像に対して過去の症例から類似した病理組織画像を検索する場合について説明する。病理組織診断とは、生体組織を顕微鏡で観察し疾病の種類を診断するものであり、腫瘍の良悪性の判別や腫瘍の種類の鑑別などのために主に行われている。

【0005】図9は従来の類似検索装置の動作を示すフローチャートである。まず、対象画像指定部1が、類似検索をおこなうべき対象画像を、検査すべき病理組織画像の中から指定する(ステップST1)。次に、特徴量抽出部2において、指定された対象画像から、その画像の特徴を定量的に表す特徴量を抽出する(ステップST2)。次に属性入力部4において、対象画像指定部1で指定された対象画像の属性を入力する(ステップST4)。対象画像の属性としては、患者名、患者ID、画像ID、腫瘍の大きさ、年齢、診断名などがあるが、診断名はこの時点では決まっていないので、入力されない。次に、特徴量抽出部2で抽出された特徴量と属性入力部4で入力された属性とをベクトル構成要素とする対象ベクトルデータを作成する(ステップST101)。図10に生成された対象ベクトルデータの一例を示す。次に、対象ベクトルデータに対して類似度が高いベクトルデータを、重みベクトル13を用いて、ベクトルデータベース7の中から類似検索エンジン10によって検索する(ステップST102)。

【0006】即ち、ベクトルデータの次元(構成要素の数)を n とし、対象ベクトルデータを $X = (x_1, x_2, \dots, x_n)$ 、ベクトルデータベースに蓄積され、上記対象ベクトルデータとの類似度を計算するベクトルデータを $Y = (y_1, y_2, \dots, y_n)$ 、重みベクトルを $W = (w_1, w_2, \dots, w_n)$ とすると、ベクトルデータ X, Y 間の類似度 $\text{sim}(X, Y)$ は以下の式で計算される。

【0007】

【数1】

$$\text{sim}(X, Y) = -\sqrt{\sum_{i=1}^n \{w_i \cdot \delta(x_i, y_i)\}^2} \quad \dots (1)$$

ここで、

$$\delta(x_i, y_i) = \begin{cases} (x_i - y_i) / (\text{属性 } i \text{ の区間幅}) & (\text{属性 } i \text{ が連続値をとる場合}) \\ 0 & (\text{属性 } i \text{ が離散値をとり、} \\ & x_i = y_i \text{ の場合}) \\ 1 & (\text{属性 } i \text{ が離散値をとり、} \\ & x_i \neq y_i \text{ の場合}) \end{cases}$$

(但し、属性 i の区間幅とは属性 i がとる最大値と最小値の差の絶対値である)

$\dots (2)$

【0008】つまり、 $\text{sim}(X, Y)$ は X と Y の重み付き距離の符号を反転したものである。このようにしてベクトルデータベース7内の全ベクトルデータを Y として対象ベクトルデータとの類似度 $\text{sim}(X, Y)$ を算出し、類似度が最大のものを選択することにより、最も類似したベクトルデータが検索される。類似度が最大のベクトルデータが複数ある場合には、それらの中のどれでも良く、例えば、最初のベクトルデータとしても良いし、ランダムに選んでも良い。

【0009】次に、検索結果表示部11において、検索された類似ベクトルデータの属性の一部と、画像データベース8内の画像のうち類似ベクトルデータに対応する画像が表示される(ステップST103)。図11に、画像と患者IDと診断名を類似度の高い順番に6つ表示する場合の表示画面例を示す。利用者は対象画像と検索の結果表示された画像を比較し、それが真に類似していると判断すれば、対象画像はその類似している画像に対応する診断名に該当する可能性が高いことがわかる。従って、対象画像に対して類似ベクトルデータと類似画像を表示することは、病理組織診断において診断名の決定を支援する上で重要な参考情報となる。次に、利用者は検索結果表示部11に表示された図11のような結果などを参考にして診断名を決定し、その診断名が類似度が最も高い類似ベクトルデータの診断名と一致すれば、利用者は類似検索の結果として回答指定部14により「正解」を、一致しなければ回答指定部14により「不正解」を指定する(ステップST104)。次に、類似検索エンジン10により最も類似するとして選出されたベクトルデータと、対象ベクトルデータと、回答指定部14により利用者により正解であるかどうかを指定された結果とをともに、重みベクトル更新部15において重みベクトル13を更新する(ステップST105)。次に、新規データ追加部12によって、対象画像の属性の

診断名の部分を利用者が決定した診断名に変更した対象ベクトルデータをベクトルデータベース7に、対象画像を画像データベース8にそれぞれ追加する(ステップST106)。

【0010】

【発明が解決しようとする課題】従来の類似検索装置は以上のように構成されているので、例えば病理組織診断に利用する場合には、類似画像として表示される画像から対象画像の正しい診断名の画像が抜け落ちる危険性があるという問題点があった。例えば、図11に示されるような類似検索結果が表示された場合、類似画像として診断名が腫瘍1、腫瘍2、腫瘍3、腫瘍7のものが含まれている。対象画像がそれらの診断名のいずれかに該当する場合には良いが、対象画像の正しい診断名が腫瘍4や腫瘍5などの類似画像として表示されなかった診断名であった場合、利用者が正しい診断名に気づかずに診断を誤ってしまう危険性があった。また、1種類の重みベクトルを使用するため、重みベクトルを最適化しても類似検索精度の大幅な向上が困難であるという問題点があった。例えば病理組織診断に利用する場合について説明すると、腫瘍1という一つの診断名に属する画像の中で類似画像を検索する際の最適な重みベクトルと、腫瘍2という別の診断名に属する画像の中で類似画像を検索する際の最適な重みベクトルは一般に異なる。これは類似度を測る上で、特徴量の個々の要素の重要性が腫瘍1と腫瘍2で通常は異なるからである。そして、腫瘍1と腫瘍2という異なる診断名を持つ画像を合わせた中から類似画像を検索する際の最適な重みベクトルは、腫瘍1に対して最適な重みベクトルと腫瘍2に対して最適な重みベクトルの中間的な重みベクトルとなる。従って、例えば腫瘍1の画像の中での類似画像検索を、その中間的な重みベクトルを用いて行った場合の類似検索精度は、腫瘍1に対して最適な重みベクトルを用いた場合よりは低

下してしまう。このように、全ての診断名を含む全画像からの類似検索を1種類の重みベクトルを用いて行う場合、重みベクトルを最適化する効果がなかなか得られにくいという問題点があった。また、対象ベクトルデータの一部となる画像の特徴量として画像から抽出された特徴量だけを用了場合、抽出の困難な特徴量が使用できず、また利用者が主観的に判断した画像の特徴を検索に利用できないという問題点があった。

【0011】この発明は上記のような問題点を解決するためになされたもので、類似検索の精度が向上でき、また重要な類似データが検索結果からぬけおちる可能性が低い類似検索装置、及び類似検索プログラムを記録した記録媒体を得ることを目的とする。

【0012】

【課題を解決するための手段】この発明の第1の構成に係る類似検索装置は、複数の対象に対して各々作成され、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象に対するベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果を個々の検索条件ごとに表示する検索結果表示部とを備えたものである。

【0013】この発明の第2の構成に係る類似検索装置は、複数の対象に対応する複数の対象画像を蓄積する画像データベースと、上記複数の対象に対して各々作成され、上記対象画像の特徴を定量的に表す特徴量と、上記対象を特徴づける属性とをベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象の画像と属性を基にベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果を対応する画像と共に個々の検索条件ごとに表示する検索結果表示部とを備えたものである。

【0014】この発明の第3の構成に係る類似検索装置は、画像の特徴量を抽出する特徴量抽出部と画像の特徴量を入力する特徴量入力部との少なくとも一方を備え、画像の特徴を定量的に表す特徴量として、上記特徴量抽出部において抽出された特徴量と上記特徴量入力部より入力された特徴量との少なくとも一方を用いたものであ

る。

【0015】この発明の第4の構成に係る類似検索装置は、画像として生体組織の画像を用い、検索条件として病理組織診断における診断名を用いたものである。

【0016】この発明の第5の構成に係る類似検索装置は、画像の特徴量として、核領域の数、核領域の面積、核領域の形状、核領域の円形度、核領域の色、核領域の色度、空隙領域の数、空隙領域の面積、空隙領域の形状、空隙領域の円形度、間質領域の数、間質領域の面積、間質領域の形状、間質領域の円形度、間質領域の色、間質領域の色度、管空領域の数、管空領域の面積、管空領域の形状、管空領域の円形度、画像のテクスチャ、ウェーブレット変換値を用いて算出される特徴量、上皮細胞が筋上皮細胞を伴う2層構造を示す程度、線維化の程度、乳頭状パターンの有無、ふんい状パターンの有無、壊死物質の有無、充実性パターンの有無、画像の色、及び色度を用いて算出される特徴量の少なくとも一つを用いたものである。

【0017】この発明の第6の構成に係る類似検索装置は、複数の検索条件にそれぞれ対応した複数の重みベクトルを蓄積する重みベクトルデータベースを備え、類似検索エンジンが用いる重みベクトルが、検索条件ごとに特定されているものである。

【0018】この発明の第1の構成に係る類似検索プログラムを記録した記録媒体は、類似検索対象に対して、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに出力する検索結果出力手順とをコンピュータに実行させるものである。

【0019】この発明の第2の構成に係る類似検索プログラムを記録した記録媒体は、類似検索対象の画像と上記対象を特徴づける属性を基に、対象画像の特徴を定量的に表す特徴量と上記対象の属性とをベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに対応する画像と共に出力する検索結果出力手順とをコンピュータに実行させるものである。

【0020】

【発明の実施の形態】

実施の形態1. 以下に本発明の実施の形態1を図に従っ

て説明する。図1は、この発明の実施の形態1による類似検索装置を示すブロック図である。図において、1は検索対象の画像を指定する対象画像指定部、2は対象画像指定部1により指定された対象画像の特徴を定量的に表す特徴量を抽出する特徴量抽出部、3は対象画像に対して利用者が認識した特徴量を入力する特徴量入力部、4は対象画像に関する特徴量以外の属性を入力する属性入力部、5は特徴量抽出部2により抽出された特徴量と、特徴量入力部3により入力された特徴量と、属性入力部4により入力された属性とをベクトル構成要素とする対象ベクトルデータを作成する対象ベクトルデータ作成部、6は複数の検索条件を生成する検索条件集合生成部、7は複数の対象に対して作成された複数のベクトルデータをデータベースとして蓄えたベクトルデータベース、8は上記複数の対象に対応する複数の画像をデータベースとして蓄えた画像データベース、9は類似検索においてベクトルデータ間の類似度を算出する際に用いる重みベクトルを蓄えた重みベクトルデータベースであり、複数の検索条件にそれぞれ対応して最適化した複数の重みベクトルを蓄積している。10は重みベクトルデータベース9に蓄えられた重みベクトルを元に、ベクトルデータベース7に蓄えられた複数のベクトルデータの中で検索条件集合生成部6により指定された属性を持つベクトルデータの中から、対象ベクトルデータに類似したベクトルデータを検索する類似検索エンジン、11は類似検索エンジン10により検索された類似するベクトルデータと、類似するベクトルデータに対応する画像とを、検索条件集合生成部6で指定された個々の検索条件ごとに表示する検索結果表示部、12は新たなベクトルデータと画像をそれぞれベクトルデータベース7と画像データベース8に登録する新規データ追加部である。

【0021】次に動作について説明する。従来技術の説明と同様、検査用の病理組織画像に対して過去の症例から類似した病理組織画像を検索する場合について説明する。図2はこの発明の実施の形態1による類似検索装置の動作を示すフローチャートである。まず、対象画像指定部1が、類似検索をおこなうべき対象画像を、検査すべき病理組織画像の中から指定する（ステップST1）。

【0022】次に、特徴量抽出部2において、指定された対象画像から、その画像の特徴を定量的に表す特徴量を抽出する（ステップST2）。特徴量としては、例えば画像中の核領域の数や面積、間質領域の数や面積などがあり、ここでは核領域の面積と間質領域の面積を抽出する。なお、これらの特徴量は例えば計算機を用いて算出される。一例として核領域の数を抽出する方法を以下に説明する。

（1）その画素が核か核でないかを判定するニューラルネットワークを作成する。

（2）そのニューラルネットワークで原画像から核だけ

を抽出した画像を作成する。

（3）ラベリングを行う。即ち、繋がっている核領域には同じラベル番号をつける。

（4）ラベルの数をカウントする。

ここで、（1）の核か核でないかを判定するニューラルネットワークの作成方法は、予めどこが核領域か判っている画像を用い、画素毎のRGB等の色情報を入力し、その画素が核かどうかの情報を教師信号として、核でないならニューラルネットワークの出力が1、核なら0となるように学習することにより行う。

【0023】次に、特徴量入力部3において、類似検索装置の利用者が画像の特徴を自ら判断し、入力する（ステップST3）。利用者が判断する特徴量としては、例えば、線維化の程度、充実性パターンの存在する程度などである。

【0024】次に、属性入力部4において、対象画像指定部1で指定された対象画像の属性を入力する（ステップST4）。対象画像の属性としては、患者名、患者ID、画像ID、腫瘍の大きさ、年齢、診断名などがあるが、診断名はこの時点では決まっていないので、入力されない。

【0025】次に、特徴量抽出部2で抽出された特徴量と、特徴量入力部3で入力された特徴量と、属性入力部4で入力された属性とをベクトル構成要素とする対象ベクトルデータが生成される（ステップST5）。図3に生成された対象ベクトルデータの一例を示す。

【0026】次に、検索条件集合生成部6において、類似検索を行う際の検索条件の集合を生成する（ステップST6）。図4に診断名を検索条件とした場合の検索条件集合の一例を示す。

【0027】次に、検索条件集合に属する個々の検索条件について、検索条件に合致し、かつ対象ベクトルデータに対して類似度が高いベクトルデータを、重みベクトルデータベース9にある重みベクトルのうち検索条件に対応した重みベクトルを用いて、ベクトルデータベース7から類似検索エンジン10によって検索する（ステップST7）。重みベクトルは個々の検索条件に対して予め重みベクトルデータベース9に登録されている最適な重みを用いる。図5に、図4に示される検索条件集合の個々の検索条件について用いられる重みベクトルの一例を示す。図4に示される検索条件集合の場合に、検索条件である診断名を満足する全画像から画像の特徴量のみを用いて類似画像を検索したい場合には、図5に示されるように、患者名、腫瘍の大きさなどの属性に関する重みは0とする。なお、重みベクトルの最適化は、ベクトルデータベースに登録されているベクトルデータを用いて、従来の類似検索装置などで使用されている重み更新手法を用いても良いし、画像データベースに登録されている画像のお互いの類似性を利用者が決め、それを評価値として類似検索結果を定量的かつ自動的に評価

できるようにしておいて、ベクトルデータベースに登録されているベクトルデータを用いて類似検索結果を評価し、その評価結果が良くなる方向へ重みを自動的に更新するようにして行っても良い。類似検索は以下のようにして行われる。ベクトルデータの次元（構成要素の数）を n とし、対象ベクトルデータを $X = (x_1, x_2, \dots, x_n)$ 、ベクトルデータベースに蓄積され、上記対象ベクトルデータとの類似度を計算するベクトルデータを $Y = (y_1, y_2, \dots, y_n)$ 、重みベクトルを $W = (w_1, w_2, \dots, w_n)$ とすると、ベクトルデータ X 、 Y 間の類似度 $\text{sim}(X, Y)$ は従来例と同様に(1)式、(2)式により計算される。ベクトルデータベース9内のベクトルデータのうち、検索条件を満足する全ベクトルデータを Y として対象ベクトルデータとの類似度 $\text{sim}(X, Y)$ を算出し、類似度が最大のものを選択することにより、最も類似したベクトルデータが検索される。類似度が最大のベクトルデータが複数ある場合には、それらの中のどれでも良く、例えば、最初のベクトルデータとしても良いし、ランダムに選んでも良い。

【0028】次に、検索結果表示部11において、検索された類似ベクトルデータの属性の一部と、画像データベース8内の画像のうち類似ベクトルデータに対応する画像が個々の検索条件ごとに表示される（ステップST8）。図6に、画像と患者IDと診断名が表示される場合の表示画面例を示す。

【0029】次に、表示画面などを参考にして利用者が診断名を決定し、新規データ追加部12によって、対象画像の属性の診断名の部分を入力した対象ベクトルデータをベクトルデータベース7に、対象画像を画像データベース8にそれぞれ追加する（ステップST9）。

【0030】以上述べたような対象画像の類似検索演算は、類似検索用プログラムにより実現され、該プログラムは記録媒体に記録して提供される。

【0031】なお、上記実施の形態では、特徴量抽出部2で抽出された特徴量はそのまま対象ベクトルデータの一部となっているが、抽出された特徴量を利用者が修正できるようにしても良い。

【0032】また、上記実施の形態では、特徴量抽出部2で抽出された特徴量と特徴量入力部3により入力された特徴量を画像の特徴量として用いているが、その一方だけを用いるようにしても良い。

【0033】また、上記実施の形態では、画像の特徴量として、核領域の面積、間質領域の面積、線維化の程度、充実性パターンが存在する程度を用いた場合を説明したが、特徴量はこれだけに限らず、この他にも核領域の数、形状、円形度、色、色度、空隙領域の数、面積、形状、円形度、間質領域の数、形状、円形度、色、色度、管空隙領域の数、面積、形状、円形度、画像のテクスチャ、ウェーブレット変換値を用いて算出される特徴

量、上皮細胞が筋上皮細胞を伴う2層構造を示す程度、乳頭状パターンの有無、壊死物質の有無、充実性パターンの有無、画像の色あるいは色度を用いて算出される特徴量などを用いても良いし、使用する特徴量は利用者が自由に設定できるようにしても良い。

【0034】また、上記実施の形態では、ベクトルデータに含める画像の属性として、患者名、患者ID、画像ID、腫瘍の大きさ、年齢、診断名を用いた場合を説明したが、この他の属性を用いても良いし、これらの属性を利用者の好みに応じて自由に設定できるようにしても良い。

【0035】また、上記実施の形態では、検索条件集合は検索条件集合生成部6において生成されているが、いつも使用するような検索条件を予め作成しておき、このような検索条件集合を指定して使用するようにしても良い。

【0036】また、上記実施の形態では、診断名を検索条件とした場合を示したが、他の属性や特徴量などを検索条件として用いても良い。

【0037】また、上記実施の形態では、予め検索条件ごとに最適化した重みベクトルを重みベクトルデータベース9に蓄えておいて使用しているが、従来の類似検索装置などで行われているように、類似検索結果が正解かどうかの利用者の判断結果などを元に重みベクトルを更新するようにしても良い。また、ある検索条件に対応する重みベクトルがなかった場合には、検索条件に近いものの重みベクトルで代用しても良いし、検索条件に近いものの重みベクトルを元に重みベクトルを新たに生成しても良いし、検索条件に対応する重みベクトルがない場合に使用すべき重みベクトルを予め決めておき、それを使用するようにしても良い。

【0038】また、上記実施の形態では、個々の検索条件に対して類似ベクトルデータが一つ検索され画像と共に表示されているが、個々の検索条件に対して類似度が高い順番に複数の類似ベクトルデータを検索し、対応する画像と共に表示するようにしても良い。

【0039】また、上記実施の形態では、(1)式と(2)式を用いて類似度を算出したが、これは一例であり、他の算出法を用いても良い。

【0040】また、上記実施の形態では、検査用の病理組織画像に対して過去の症例から類似した病理組織画像を検索する場合の一例について説明したが、細胞診の画像やX線写真を初めとした他の医用画像に関する類似検索やりんごなどの農産物の等級の判別など、他の様々な画像を用いた類似検索においても同様の効果を奏する。また、必ずしも画像を用いる必要はなく、例えば医療上の診療データを用いた過去の類似症例の検索など、一般にベクトルデータの類似検索において同様の効果を奏する。

【0041】実施の形態2. 図7は実施の形態2による

類似検索装置を示すブロック図であり、画像を用いない類似検索の例である。ここでは患者の診療データから過去の類似した診療データを検索する場合について説明する。図において、属性入力部4には対象を特徴づける複数の属性として、例えば、患者名、患者ID、年齢、職業、熱、せきの有無、鼻水の有無、のど痛の有無、診断名等が入力され、上記複数の属性を、ベクトル構成要素とするベクトルデータが対象ベクトルデータ作成部5で作成される。なお、診断名はこの時点では決まっていないので、入力されない。検索条件集合生成部6には例えば風邪や花粉症等の診断名が検索条件として複数入力される。類似検索エンジン10では、検索条件集合生成部6に入力された検索条件の個々の検索条件について、検索条件に合致し、かつ対象ベクトルデータに対して類似度が高いベクトルデータを、重みベクトルデータベース9にある重みベクトルのうち検索条件に対応した重みベクトルを用いて、ベクトルデータベース7の中から検索する。ここで重みベクトルは、実施の形態1と同様、個々の検索条件に対して予め重みベクトルデータベース9に登録されている最適な重みを用いる。検索結果は検索結果表示部11に表示される。

【0042】

【発明の効果】以上のように、この発明の第1の構成によれば、複数の対象に対して各々作成され、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象に対するベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果を個々の検索条件ごとに表示する検索結果表示部とを備えたので、利用者が参照したい個々の検索条件に対応した類似データが検索され表示され、重要な類似データが検索結果から漏れる可能性が低くなるという効果がある。

【0043】また、この発明の第2の構成によれば、複数の対象に対応する複数の対象画像を蓄積する画像データベースと、上記複数の対象に対して各々作成され、上記対象画像の特徴を定量的に表す特徴量と、上記対象を特徴づける属性とをベクトル構成要素とするベクトルデータを複数蓄積するベクトルデータベースと、指定された類似検索対象の画像と属性を基にベクトルデータを作成する対象ベクトルデータ作成部と、複数の検索条件を生成する検索条件集合生成部と、この検索条件集合生成部で生成された個々の検索条件ごとに、上記ベクトルデータベースに蓄積された複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデー

タに類似するベクトルデータを検索する類似検索エンジンと、上記類似検索エンジンにより検索された結果に対応する画像と共に個々の検索条件ごとに表示する検索結果表示部とを備えたので、画像を伴った重要な類似データが検索結果から漏れる可能性が低くなるという効果がある。また、類似したベクトルデータと共に類似画像を表示できる効果がある。

【0044】また、この発明の第3の構成によれば、画像の特徴量を抽出する特徴量抽出部と画像の特徴量をを入力する特徴量入力部との少なくとも一方を備え、画像の特徴を定量的に表す特徴量として、上記特徴量抽出部において抽出された特徴量と上記特徴量入力部より入力された特徴量との少なくとも一方を用いたので、画像の特徴量として、画像から抽出された特徴量の他に利用者が自ら判断した主観的な特徴量を使用でき、抽出の困難な特徴量が利用できたり、利用者の好みに応じた類似検索ができる効果がある。

【0045】また、この発明の第4の構成によれば、画像として生体組織の画像を用い、検索条件として病理組織診断における診断名を用いたので、病理組織診断を有効に支援できる効果がある。特に、稀な診断名に対しても、その診断名に属する画像から対象画像に類似する画像が少なくとも一画像表示されるため、稀な診断名を見落とす危険性が低くなる効果がある。また、画像が紛らわしい診断名の場合も、紛らわしい診断名全てについて類似画像を表示できるため、それらを比較検討することにより、紛らわしい場合に診断を誤る危険性を低減できる効果がある。

【0046】また、この発明の第5の構成によれば、画像の特徴量として、核領域の数、核領域の面積、核領域の形状、核領域の円形度、核領域の色、核領域の色度、空隙領域の数、空隙領域の面積、空隙領域の形状、空隙領域の円形度、間質領域の数、間質領域の面積、間質領域の形状、間質領域の円形度、間質領域の色、間質領域の色度、管空隙領域の数、管空隙領域の面積、管空隙領域の形状、管空隙領域の円形度、画像のテクスチャ、ウェーブレット変換値を用いて算出される特徴量、上皮細胞が筋上皮細胞を伴う2層構造を示す程度、線維化の程度、乳頭状パターンの有無、ふるい状パターンの有無、壊死物質の有無、充実性パターンの有無、画像の色、及び色度を用いて算出される特徴量の少なくとも一つを用いたので、病理組織診断における検索精度をより向上できる効果がある。

【0047】また、この発明の第6の構成によれば、複数の検索条件にそれぞれ対応した複数の重みベクトルを蓄積する重みベクトルデータベースを備え、類似検索エンジンが用いる重みベクトルが、検索条件ごとに特定されているので、類似検索精度をさらに向上できる効果がある。

【0048】また、この発明の第1の構成に係る類似検

索プログラムを記録した記録媒体によれば、類似検索対象に対して、対象を特徴づける複数の属性をベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに出力する検索結果出力手順とをコンピュータに実行させるようにしたので、重要な類似データが検索結果から漏れる可能性が低くなるという効果がある。

【0049】また、この発明の第2の構成に係る類似検索プログラムを記録した記録媒体によれば、類似検索対象の画像と上記対象を特徴づける属性を基に、対象画像の特徴を定量的に表す特徴量と上記対象の属性とをベクトル構成要素とするベクトルデータを作成する対象ベクトルデータ作成手順と、複数の検索条件に対して、個々の検索条件ごとに、ベクトルデータベースに蓄積された複数の対象に対する複数のベクトルデータの中から、上記検索条件を満足し、かつ上記対象ベクトルデータに類似するベクトルデータを検索する類似検索手順と、検索された結果を個々の検索条件ごとに対応する画像と共に出力する検索結果出力手順とをコンピュータに実行させるようにしたので、画像を伴った重要な類似データが検索結果から漏れる可能性が低くなるという効果がある。また、類似したベクトルデータと共に類似画像を表示できる効果がある。

【図面の簡単な説明】

【図1】 この発明の実施の形態1による類似検索装置

を示すブロック図である。

【図2】 この発明の実施の形態1による類似検索装置の動作を示すフローチャートである。

【図3】 この発明の実施の形態1に係わる対象ベクトルデータの一例である。

【図4】 この発明の実施の形態1に係わる検索条件集合の一例である。

【図5】 この発明の実施の形態1に係わる重みベクトルの一例である。

【図6】 この発明の実施の形態1による類似検索装置で表示される表示画面の一例である。

【図7】 この発明の実施の形態2による類似検索装置を示すブロック図である。

【図8】 従来の類似検索装置を示すブロック図である。

【図9】 従来の類似検索装置の動作を示すフローチャートである。

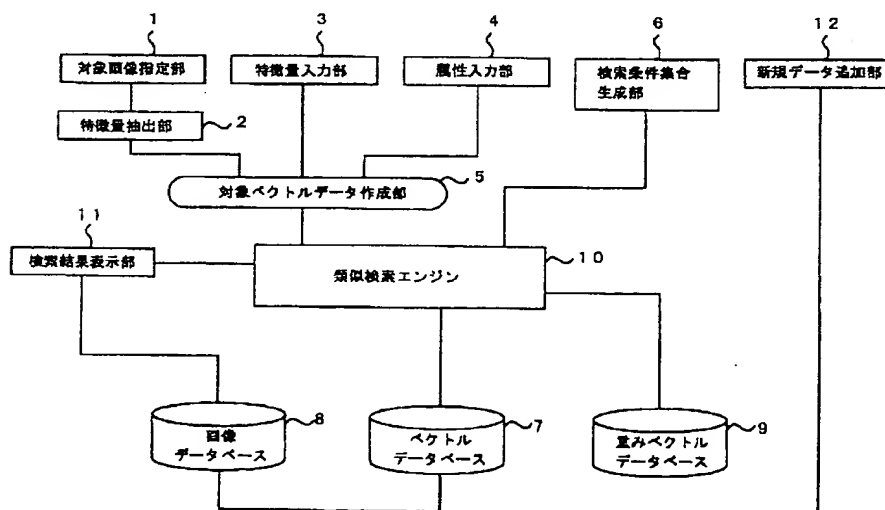
【図10】 従来の類似検索装置で用いる対象ベクトルデータの一例である。

【図11】 従来の類似検索装置で表示される表示画面の一例である。

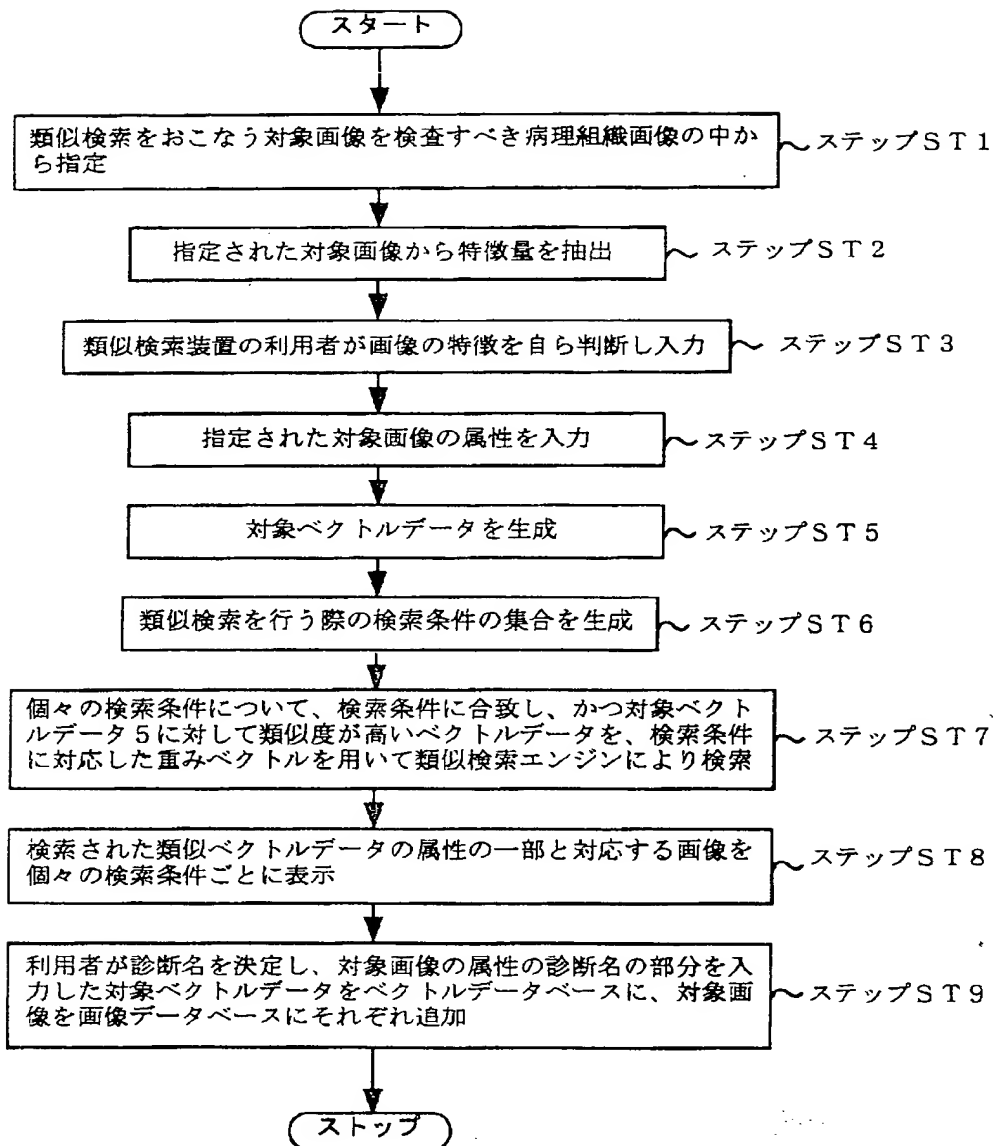
【符号の説明】

1 対象画像指定部、2 特徴量抽出部、3 特徴量入力部、4 属性入力部、5 対象ベクトルデータ作成部、6 検索条件集合生成部、7 ベクトルデータベース、8 画像データベース、9 重みベクトルデータベース、10 類似検索エンジン、11 検索結果表示部、12 新規データ追加部、13 重みベクトル、14 回答指定部、15 重みベクトル更新部。

【図1】



【図2】



【図3】

患者名	患者ID	画像ID	腫瘍の大きさ	年齢	診断名	核領域の面積	間質領域の面積	組織化の程度	充実性パターンの存在する程度
三浦花子	970088	970123	2.5	35	-	0.12	0.74	0.2	0.05

【図4】

検索条件名	検索条件
検索条件1	診断名＝腰痛1
検索条件2	診断名＝腰痛2
検索条件3	診断名＝腰痛3
検索条件4	診断名＝腰痛4または腰痛5
検索条件5	診断名＝腰痛6
検索条件6	診断名＝腰痛7または腰痛8

【図10】

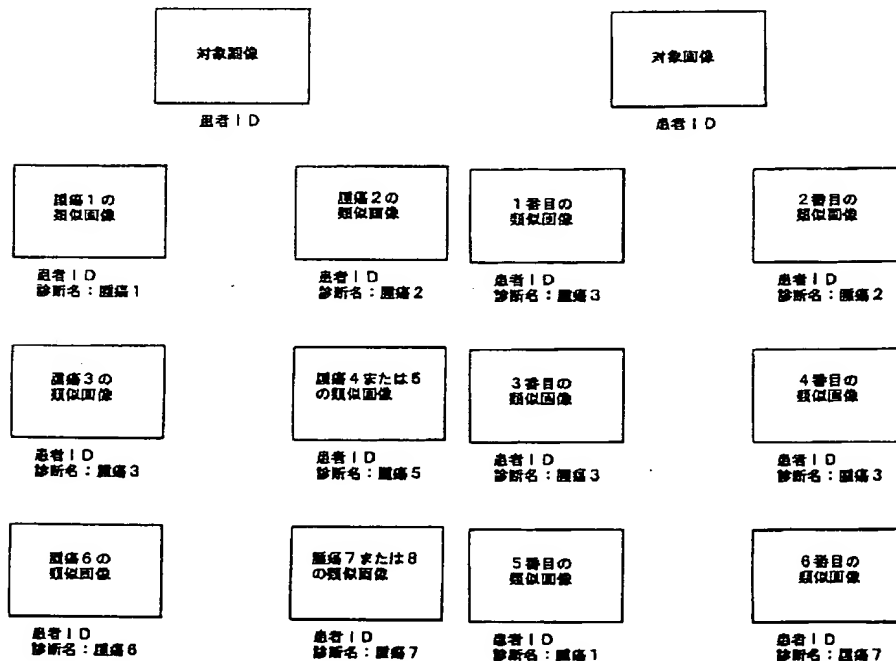
患者名	患者ID	画像ID	腰痛の 大きさ	年齢	診断名	特徴量1	特徴量2	特徴量3	特徴量4
三夏 花子	970088	970123	2.5	35	—	0.05	0.12	0.74	0.2

【図5】

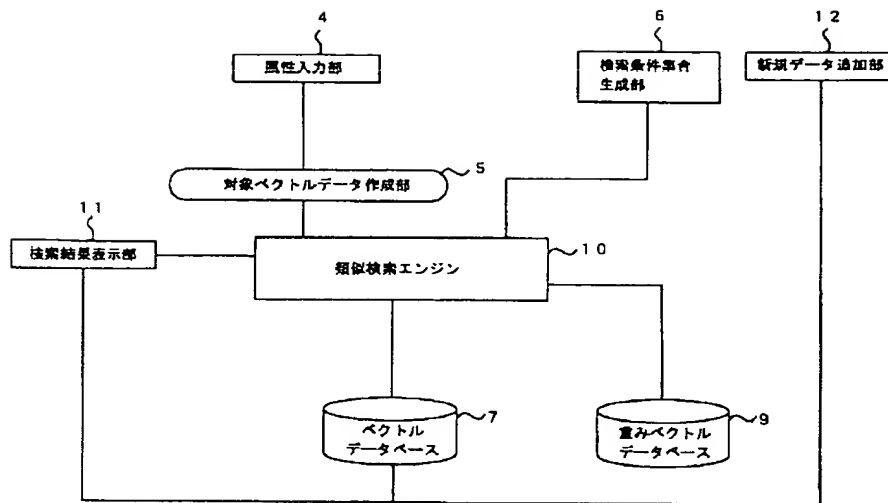
	患者名	患者ID	画像ID	腰痛の 大きさ	年齢	診断名	椎間域 の面積	間質領域 の面積	繊維化 の程度	充塞性パターンの 存在する程度
検索条件1	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.2	0.5	0.1
検索条件2	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3	0.0	0.0
検索条件3	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.2	0.8
検索条件4	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.7	0.0	1.0
検索条件5	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.8	0.8	0.2
検索条件6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.5

【図6】

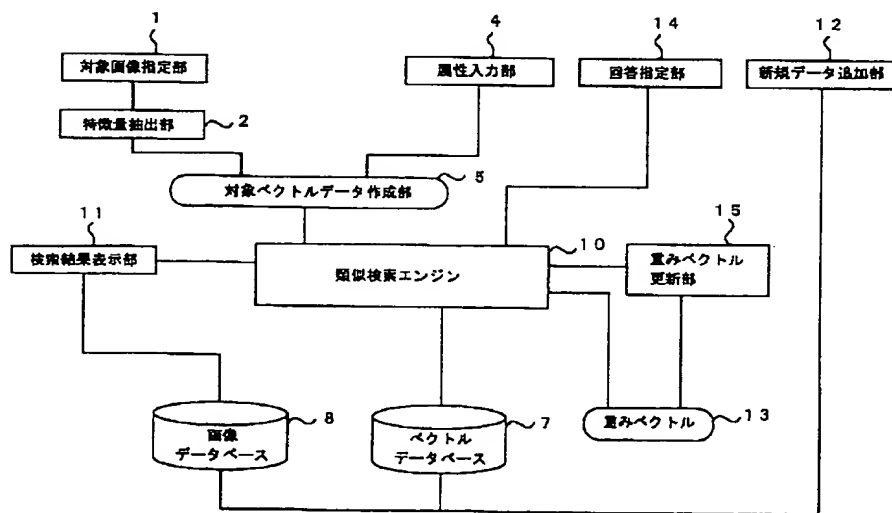
【図11】



【図7】



【図8】



【図9】

